We sincerely thank the reviewers for providing thoughtful comments on our manuscript. We have provided a point-by-point response to these comments below. We are confident that these additions and revisions improve the content of our manuscript for readers of Science of the Total Environment.

## Reviewer 1

This manuscript is a welcome addition in the world of trend analyses. The authors propose combinations of GAM and mixed-effects meta-analysis regression to account for uncertainty. The methodology is sound and results are convincing. My recommendation would be even more enthusiastic if the authors had formatted their methods sections in a more traditional manner. I am of the opinion that models should be presented using clear mathematical equations and not in R code semantics. This comments and a few other can be found in the attached document.

* **Response**: We really appreciate these comments that affirm our opinion on the value of this work for trend analysis of long-term water quality data. We have changed our formulas and text to a conventional format, as opposed to R code notation. Briefly, here are the updated equations for the models described in the methods.

## Reviewer 2

### Overview

This manuscript presents a new application of GAMs approaches to fit irregular time series data (gaps, changing frequency of measurements) to estimate trends. The paper presents a case study using Chl a data from San Francisco Bay. The manuscript is presented as a general approach with applications to other similar environmental time series data. A comparison of methods shows that the GAMs plus meta-analysis method provides some different trend conclusions from ordinary least squares and plain GAMs. The paper is generally well written and well-suited for the audience for STOTEN. I have some clarifications that would help improve/clarify the utility and readability of the manuscript. These are presented below.

* **Response**:

### Utility (1):

We applied the wqtrends code in R to a time series data set of X2 (salinity intrusion measure) over approximately in San Francisco Bay. Some observations:

-the code did not work with the full time series of daily X2 data. But we were able to make it work using a random sampling of 5 points per month. Could the reason for this error be explored? As time series data are concerned, we provided an input of ~100 years of daily data. A sample result for the trends are shown in the figure below. I think it would be helpful for readers/users to understand what the practical limitations of the code/algorithm are.

* **Response**:

### Utility (2):

From figure 7 it seems that there are very different conclusions for Station 36 depending on the method used. From Figure 8, however, it seems that Station 36 is a bit of an anomaly in this regard, and that other 8 stations in the case study have findings on trends that are similar among the three methods—and that the results are not dramatically different. Perhaps this should be the basis for added discussion in the final section—when is the added complexity of the GAMS and meta analysis method is appropriate in a real-world setting? Alternatively, it would be helpful if the authors were to examine another data set from the region that they work with, to review the relative findings of trend significance and what this depends on.

* **Response**:

### Readability/Clarity:

The term meta-analysis seems to be an important part of the paper and is referred to multiple times, but it is never clearly explained in the manuscript. For example, at line 164. See the following sentences.

“Third, we used a mixed-effects meta-analysis to estimate trends and test hypotheses about the change in seasonal averages across years. While 166 meta-analysis methods arose from analyses of results from multiple studies, their distinguishing characteristic is propagation of uncertainty (Gasparrini et al., 2012; Sera et al., 2019). Meta-analysis uses response data that includes standard errors (uncertainties) as needed to address our questions.”

* **Response**:

It is not clear to me what exactly is meant by meta-analysis and what is being tested in the context of the time series data. An explanation that is understandable by a general audience of potential users of this method would be helpful. It does seem likely, as that sentence points out, that most readers would be more familiar with meta-analysis methods in the context of aggregating the results of multiple studies. There are other helpful citations afterwards, but a more explicit connection between the familiar use case and the use case here of combining different seasonal features from a single regression model would have been helpful for reading this work.

* **Response**: